

Examining the Yahoo! Sponsored Search Auctions: A Regression Discontinuity Design Approach

Jia Yuan

Faculty of Business Administration, University of Macau

L212, FBA, University of Macau, Macau

Tel: 853-8397-4162 E-mail: jiayuan@umac.mo

Received: December 3, 2011

Accepted: January 17, 2012

Published: March 1, 2012

doi:10.5539/ijef.v4n3p139

URL: <http://dx.doi.org/10.5539/ijef.v4n3p139>

Abstract

The sponsored search auction is a successful pricing mechanism which helps search engine companies sell navigation service to advertisers and generate multibillion dollar revenue. For the two popular sponsored search auctions—the Generalized First Price (GFP) auction and the Generalized Second Price (GSP) auction—current consensus in both the industry and academia is that the GSP auction is more stable than the GFP auction. Specifically, in the GSP auction, bidders are less likely to “game the system”, meaning that an individual bidder will change his bid less frequently and his bid range will be smaller. This paper examines this prevailing belief using a Regression Discontinuity Design (RDD) approach and finds that after bidders switch to the GSP auction, they actually change their bid 19% more frequently and increase their daily bid range by 46%. The paper suggests that the prevailing automated bidding strategies should not be ignored in explaining bidders’ behavior.

Keywords: Keyword auction, Sponsored search, Yahoo! Business Model, Regression discontinuity

1. Introduction

The sponsored search auction has played an indispensable role in the success of search engine giants like Yahoo! and Google. For example, Yahoo!’s first half-year revenue in 2008 was \$3.62 billion, and at least 50% of that revenue came from the sponsored search auction. For Google, its first half-year revenue in 2008 was \$10.55 billion, with 97% of this revenue generated by the sponsored search auctions. Actually, the sponsored search auction is not only crucial to search engine companies, but it is also “vital to the success of many other small businesses” such as bid management software firms, bidding campaign consulting firms, and key word selecting firms, etc. (See Jansen and Mullen, 2008).

The sponsored search auction is a pricing mechanism which helps search engine companies sell navigation services to advertisers. When addressing search requests, search engines display both the search results and advertisers’ web links, which are called sponsored links. These sponsored links attempt to navigate potential customers to specific product web sites. Because this targeting of potential customers has proven effective, advertisers are willing to pay in order to obtain an ideal placement for their web link on a search result page. Search engine companies invented the sponsored search auction to sell these sponsored link placements.

The sponsored search auction was first introduced in 1998 by Goto for Yahoo!. Since then, search engine designers have upgraded the mechanism several times. The purpose of replacing an old sponsored search auction with a new one is “to bring more stability to the auction bidding, increase profits, and help reduce strategic bidding” (Jansen and Mullen, 2008). For search engines, the best scenario is that each of the bidders truthfully spends his maximum willingness to pay on purchasing the links. However, as pointed out by Edelman et al. (2007), bidders have incentive to game the auction system to reduce their costs. In theory, McAdams and Schwarz (2007) argue that the costs that buyers incur while trying to “game” an auction mechanism are fully passed through to the seller. Therefore, search engines may be sensitive to the stability of their sponsored search auction mechanisms.

Correctly understanding how sponsored search auctions work will not only affect the multibillion dollar revenue of search engine companies, but it will also help develop superior sponsored search auctions in the future. In their seminal and influential work, Edelman et al. (2007) and Varian (2007) propose a Generalized Second Price (GSP) auction framework to understand the sponsored search auctions from 2002 through 2007. Edelman et al. (2007) and Varian (2007) argue that this game framework “describes the basic properties of the prices observed in Google’s ad auction reasonably accurately.” However, Borgers et al. (2007) suggest that “this static GSP auction model actually

may have a very poor explanation power on the real data collected from the Yahoo! sponsored search auction". Actually, the GSP auction approach ignores many important aspects of the actual bidding, such as the dynamic bidding nature or the fact that, in practice, bidders often employ automated bidding software to manage their bidding. Therefore, to what extent this GSP framework can explain the actual bidding behaviors still needs careful empirical examination.

This paper reinvestigates this issue by exploiting a natural experiment which took place on June 26, 2002. On that day, Yahoo! switched its sponsored search auction from the Generalized First Price (GFP) auction to the Generalized Second Price (GSP) auction. Auction theory claims that "superior designs" have replaced the "inefficient market institutions" (Edelman et al., 2007 and Jansen and Mullen, 2008). The GSP game framework predicts that, in the new auction, bidders will be less likely to "game the system". This means that an individual bidder will change his bid less frequently and that his bid range will be smaller. This paper tests these hypotheses using bid data collected from Yahoo! sponsored search auctions in 812 keyword markets from June 15, 2002 through June 14, 2003.

The econometric method adopted in the test is the Regression Discontinuity Design (RDD) approach. Yahoo!'s auction rule change on June 26, 2002 provides a treatment effect framework. Specifically, all the bidders after June 26, 2002 would face the treatment of the GSP auction. Thus, estimating the bidding behavior difference in these two auction systems will be turned into identifying the average treatment effect. However, in the standard treatment effect framework, identification usually depends on strong assumptions. For instance, for the identification of OLS regression, one has to assume that the error terms of the regression satisfy certain ad hoc statistical relationships such as the i.i.d. assumption. This paper avoids this challenge by using a Regression Discontinuity Design (RDD) approach, which enables us to make relatively simple and reasonable assumptions to obtain identification.

The estimation results show that the bidding behavior under the new auction system was less stable than the GSP auction framework predicted. The daily frequency with which an individual bidder changed his bid increased by 3.7 times, representing a 19% increase. In addition, the daily bid range of each bidder increased by \$0.71, representing a 46% increase. All the above findings contradict the prediction made by the GSP auction framework which stated that in the new auction bidding behaviors would be more stable.

These findings are important in two aspects. First, in practice, the stability of the sponsored search auction system is an important property. In theory, McAdams and Schwarz (2007) argue that the costs that buyers incur while trying to "game" an auction mechanism are fully passed through to the seller. Therefore, search engines may be sensitive to the stability of their sponsored search auction mechanisms. Actually, in practice, stability has become an important concern. One of the reasons why Yahoo! switched to the GSP auction system was that Yahoo! believed that the GSP auction was more stable and the "second-price structure makes the market less susceptible to gaming" (Edelman et al., 2007).

Second, the findings above have important implications for the current sponsored search auction theory literature. Recent theories on the sponsored search auction, including Edelman et al. (2007), Varian (2007) and Athey and Ellison (2007), are basically based on a static game theory structure. One salient fact these theories miss is that bidders often adopt automated bidding robots to manage their bidding. Yuan (2009) suggests that over 40% of bids may be submitted by these bidding robots following certain technical bidding rules, such as Position Targeting, Cost-per-Purchase Bidding, ROI Bidding, and so on.

It should be noted that this paper is not the first empirical work which provides evidence and suggests that the GSP auction framework may have trouble explaining the actual bidding data. Borgers et al. (2007) suggest that "this static GSP auction model actually may have a very poor explanation power on the real data collected from the Yahoo! sponsored search auction". The difference between this paper and Borgers et al. (2007) is that the latter adopt a structural estimation approach. Plus, to obtain identification, they have to make relatively strong assumptions. Besides, their data set is relatively small. However, this paper exploits the institutional variation and conducts a reduced form regression by employing a RDD approach. This approach enables the author to make relatively simple and clean assumptions. Another advantage this paper has is the huge data set. This paper uses over 1 million bids collected from 812 keyword markets to conduct the RDD estimation, as opposed to the roughly 60,000 bids collected from 5 keyword searches used by Borgers et al. (2007).

The paper is organized as follows. Section 2 introduces the Yahoo! sponsored search auction. Section 3 briefly surveys the sponsored search auction theory literature and proposes the hypothesis to test. Section 4 introduces the data and presents the simple statistics and OLS regression results. Section 5 sets up the RDD framework for the test and discusses the estimation methods. Section 6 presents the RDD estimation results. Section 7 concludes.

2. Yahoo! Sponsored Search Auction

In the search engine industry, there are three key players: the advertisers, the search engines, and the potential customers. Search engines navigate potential customers to advertisers' product web sites by displaying their web

links when potential customers conduct keyword search requests. Advertisers' links are called sponsored links. Sponsored links distinguish themselves from the organic (non-sponsored) web search results by whether or not a fee is paid to the search engine company.

Figure 1 shows an example of sponsored links for the key word "refinance". When someone uses Yahoo! to search for information about "refinance", the search engine will display search results along with sponsored links, which are circled in Figure 1. Usually around 10 sponsored links, located on the top and on the right of each page, will be displayed. Advertisers are interested in buying these link slots for their product web sites because they may target the potential customers more efficiently. In 1998, Goto first introduced the sponsored search auction in the search engine industry to sell these link slots.

The sponsored search auction is a multi-object dynamic auction in which each individual advertiser bids for the ideal slot for his web site. Sponsored search auctions usually have the following common features. First, all the link slots are auctioned at the same time. As shown in Figure 1, there were at least 12 sponsored link slots being auctioned at that time. Second, the auction is dynamic with an infinite time horizon. Each bidder can change or withdraw his bid at any time, which will be immediately reflected in his slot placement. Third, all search engines share a common payment rule: pay per click (PPC), which means that whenever there is a click on the sponsored link, the bidder will pay Yahoo! once. And lastly, in Yahoo!'s sponsored search auction, all the information, including bids and slot placement, is public information, which can be observed by all the bidders directly.

In keeping with the keyword search for "Refinance", Figure 2 shows all bidders' bids and slot allocation information as it was captured by a free public web site. This free bid check website is http://keyword.secretstohighpro_t.com/default.aspx. Actually, Figure 1 and Figure 2 were captured at the same time on March 28, 2007. Figure 2 shows that the bid range is from \$16.13 to \$7.49 and each bidder's position is determined solely by his bid. As can be seen, "Lending Tree" had the highest bid; therefore, this advertisement was placed at the highest slot as shown in Figure 1.

Designing efficient auction rules regarding how advertisers pay a search engine and how a search engine allocates the link slots among its advertisers is a key challenge faced by search engine designers because the decision to adopt different forms of sponsored search auctions has an important impact on the success of search engine companies. In the past six years, Yahoo! upgraded its sponsored search auction several times hoping to find a better auction mechanism to bring more stable bidding behaviors and higher auction revenue.

Before June 26, 2002, a bidder in the Yahoo! sponsored search auction paid Yahoo! his bid multiplied by the number of the clicks on his web site. For example, if a bidder bid \$3 and his web site received 3,000 clicks, the bidder would have to pay Yahoo! \$9,000. The literature calls this type of sponsored search auction "Generalized First Price (GFP) Auction" to distinguish it from the standard first price auction.

On June 26, 2002, Yahoo! upgraded its Generalized First Price (GFP) Auction to a Generalized Second Price (GSP) Auction. In this new auction system, the web site placement was still determined solely by a bidder's bid, but each bidder, instead of paying his own bid per click, only had to pay 0.01 more than the next highest bid below his. For example, if two bidders bid \$0.4 and \$0.6, respectively, in the old bidding system, the winner would pay \$0.6 per click received; however, in the GSP auction system, he would be charged at a rate of \$0.41.

3. Literature

Recently, economists have been interested in providing a theoretical game foundation for the sponsored search auction between 2002 and 2007. Edelman et al. (2007) first introduced equilibrium concepts for the GSP auctions based on the idea of "envy-free", which assumes that in the equilibrium no bidder would like to place a bid that would cause retaliation. Varian (2007) independently discovered "envy-free" equilibrium and named it "Symmetric Nash Equilibrium" (See Edelman et al. 2007). Many papers follow these two papers and expand the theory framework. In a similar setup, Athey and Ellison (2007) further introduce consumer search behavior into the model and analyze the implications for reserve prices, product variety, and etc. Borgers et al. (2007) make more general assumptions on bidding behaviors and find the multiplicity of Nash Equilibrium.

Edelman et al. (2007) and Varian (2007) argue that this game framework "describes the basic properties of the prices observed in Google's ad auction reasonably accurately," Yet Borgers et al. (2007) suggest that "it actually may have a very poor explanation power on the real data collected from the Yahoo! sponsored search auction." This debate suggests that we still need to carefully conduct empirical examinations on the issue.

In the following subsection, this paper will examine the current GSP auction theory and construct testable hypotheses.

3.1 The Prediction of the GSP Auction Framework

If the current static GSP auction accurately describes bidders' bidding behavior, the actual bidding should follow the

predictions of the theory when it comes to the comparison between the GSP auction and the GFP auction. Currently, theories mainly focus on the GSP auction. In contrast, hardly any formal theoretical analysis has been done on the GFP auction. To compare these two auctions, Edelman et al. (2007) proposed a simple but influential example, which the following literature frequently cited (See Edelman and Ostrovsky, 2006; Jansen and Mullen, 2008). In this subsection, this research also follows this example and constructs the hypothesis.

Example 1 (Edelman et al., 2007): *There are two slots for the links. The first slot receives 400 clicks per hour, and the second slot receives 100 clicks per hour. There are three advertisers bidding to place their product. The value per click for the bidders is \$5, \$4 and \$2. Call these three bidders A, B, and C respectively.*

Edelman et al. (2007) use this example to illustrate the superiority of the GSP auction over the GFP auction in terms of stability. They show that in the GSP auction, the equilibrium bids of A, B, and C will be \$5, \$4 and \$2 and that with these bids, efficient allocation is achieved. In the GFP auction, however, the equilibrium will not be stable. B will bid \$2.01 instead of \$4 and A will bid \$2.02 instead of \$5. B will outbid A at \$2.03 and the bids will escalate until \$4. B will pull his bid back to \$2.01 and the bid escalation will go on again. These bidding behaviors will result in the sawtooth pattern of a bidding war, which is well documented in the literature (Edelman and Ostrovsky, 2006) and (Zhang, 2005). Based on this example, they argue that this second price auction structure makes the GSP auction more stable than the GFP auction.

In the following research, this paper will test the above prediction by estimating how much more stable are the bidding behaviors in the GSP auction. To be specific, if the static GSP auction framework works as the theories claim, the following hypotheses can be tested:

Hypothesis 1 On average, an individual bidder in the GSP auction will change his bid less frequently than in the GFP auction.

Hypothesis 2 On average, in the GSP auction an individual bidder's bid range will be smaller than in the GFP auction.

4. Data

Yahoo!'s research department provides a data set which records all of the bids for the top 812 keyword search by volume and all of the associated accounts for the time period from June 15, 2002 through June 14, 2003.

Each observation in the data contains the following information: bidder ID, his bid, the time when the bid was submitted, and the keyword ID. Both advertisers and keyword phrases are represented as meaningless anonymous numbers so that no identifying information is revealed. The system only records the bid change for 15 minute increments. Therefore, within the 15 minute window, any ordering information cannot be interpreted as chronological.

Table 1 shows several market statistics: the max bid, the mean bid, the minimum bid, and the standard deviation for the top ten most clicked markets. Five cents is the minimum requirement for bidding. One striking observation is the value of the maximum bid. According to this data set, one bidder is paying Yahoo! \$9,170 for just one click through the sponsored search.

Table 2 presents the possible evidence of automated bidding behaviors. The first row of the table shows the simple statistics of the number of times each bidder changed his bid each day. On average, among all bidders and all markets, each bidder changed his bid 15 times each day. The maximum number of times a bid changed in one day is 17,867, which is definitely the result of automated bidding. The lower part of the table gives some ideas about the portion of bids submitted by automated bidding software. First, it shows that more than 53% of bidders update bids 40 times each day. In other words, if updating bids 40 times each day implies the employment of automated bidding, around 53% of bids are submitted by bidding robots. Moreover, the table shows that more than 44% of bidders update bids 100 times each day and more than 28% of bidders change their bids over 500 times each day, which is probably the case of automated bidding. To summarize, Table 2 reveals that a significant portion of bids are submitted through automated bidding software.

The prevailing employment of automated bidding software has important implications because the software usually follows certain mechanical rules instead of solving optimization problems. This will raise questions about how accurately the current theories can describe the actual bidding behaviors in the sponsored search auctions.

I also present the individual bidding statistics from June 15, 2002 through July 15, 2002 in Table 3. Table 3 provides the maximum value, the mean value, the minimum value, and the standard deviation for the following daily statistics:

Bid frequency: the number of times that an individual bidder changes his bid each day.

Bid range: the difference between the maximum bid and the minimum bid of each bidder on each day.

The maximum bid, 75th percentile bid, mean bid, median bid, and 25th percentile bid of each bidder on each day.

Bid frequency and bid range measure the bidding stability of the auction system. The maximum bid, 75th percentile bid, median bid, mean bid and 25th percentile bid measure the impact on the bid distribution of an individual bidder.

Table 3 shows how the mean values of the above statistics change after the launch of the new auction. The mean values of both the daily bid frequency and the daily bid range increase, which suggests that the new auction system is more unstable. The mean values of the max bid and the 75th percentile bid increase while the mean value of the 25th percentile bid decreases, which suggests that the bids are more dispersed.

It is impossible to plot the statistics because of the huge size of the data. To further show a bigger picture of how the change of auction systems affects bidding behaviors, we run a simple OLS regression. In the OLS regression, we control for the market fixed effect and weekday effect. We also control for the entry and exit of bidders by focusing on bidders who submit bids both before and after the auction rule change.

The OLS regression examines the percentage change of the variable $y_{i,m,t}$ in the following three months from July 2002 to September 2002.

$$\log(y_{i,m,t}) = \mu_m + \sum_{j=1}^3 \alpha_j \cdot I_j(t) + \sum_{d \in \{T,W,Th,F,Sa,Su\}} \beta_d \cdot I(t=d) + u_{i,m,t}$$

$y_{i,m,t}$ is the statistic of interest for bidder i , market m , at time t . μ is the market fixed effect. $I_j(t)$ is an indicator function, showing whether the time t is in the j^{th} month after the policy change. Therefore, α_1 measures the impact of the new auction system on the market in the first month after the policy change. α_2 measures the impact in the second month after the policy change, and so on. β_d is the weekday dummy for Tuesday through Sunday.

Table 4 provides the estimation results, showing the bidding behavior change after June 26, 2002. First, both the daily bidding frequency and the daily bid range increase after the auction rule change. This result shows that the bidding behavior is more unstable under the GSP auction in contrast to the predictions which suggest the opposite result.

Second, the individual daily bid distribution expands as the percentage changes of the daily maximum bid, mean bid and median bid are bigger than that of the daily 25th percentile bid.

It may be too ambitious to claim the rejection of the hypotheses based only on the OLS regression results. The unobserved heterogeneities will make these OLS estimations biased, even misleading. First, the OLS regression does not control for the competition brought by the entry and exit of bidders in each market, which is also impossible to do. Fewer numbers of bidders might make the bidding behavior less aggressive. This may bring bias to the estimation of the average daily bidding frequency and the daily bidding range.

Second, the OLS regression cannot control for the competition among search engine markets. During that time period, around 2002, Google's popularity was beginning to rise, becoming more and more popular and grabbing more and more sponsored search market shares. Bidders often had accounts in both search engines. The implication, thus, would be that bidders might have been transferring more resources to Google accounts and paying less attention to their Yahoo! searches. This might also have caused bidders to bid less aggressively, which would affect our OLS estimation.

Moreover, the OLS regression cannot control for many unobserved variables which play an important role in the bidding strategies, such as the bidders' budgets, the conversion rate of purchases, and so on.

The bottom line here is that although the above regressions present a bigger picture of the bidding behavior change and suggest that the GSP auction does not bring stability to the system, the story is not convincing as there are so many unobserved variables which might plague the estimation. Therefore, to identify the true average effect, at least to reveal the sign of it, in the following section we use a RDD approach to avoid the above impacts of the unobserved heterogeneities.

5. Model

The basic idea behind the estimation is to exploit the natural experiment to identify the treatment effect of the GSP auction, which is the performance difference between the two auction systems.

Let $y_i(x_i, t)$ denote the statistics of interest for individual i at time t . X_i is a variable vector including all other

characteristics such as market dummy and weekday dummy.

After June 26, 2002, Yahoo! launched a new auction system. Let \bar{t} be the critical time when this new auction started. Then the statistics y_i can further be rewritten as

$$y_i(x_i, t) = \beta_i(x_i, t) + \alpha_i(x_i, t) \cdot I_i(t \geq \bar{t}) + u_i(x_i, t) \tag{1}$$

and $\alpha_i(x_i, t)$ captures the impact the new auction rule has on bidders' bidding behavior. $I_i(t \geq \bar{t})$ is the indicator function of whether the bids were submitted after \bar{t} . There is no restriction on the form of $\beta_i(x_i, t)$ except continuity at \bar{t} . $E[\alpha_i(x_i, t)]$ is the average treatment effect, which is what we want to estimate.

If the hypothesis is valid, $E[\alpha_i(x_i, t)]$ should be negative. The challenge is the identification of $E[\alpha_i(x_i, t)]$ as discussed in section 4. The unobserved heterogeneities will make OLS estimations biased, even misleading. However, the RDD approach can obtain identification with more relatively simple and clean assumptions.

5.1 The Identification by RDD

Following Imbens and Angrist (1994) and Cameron and Trivedi (2005), it is easy to show the following:

$$E[\alpha_i(x_i, t)|t = \bar{t}] = \lim_{t \downarrow \bar{t}} E[y_i(x_i, t)|t] - \lim_{t \uparrow \bar{t}} E[y_i(x_i, t)|t]$$

RDD approach assumes that the bidding statistics under the GFP auction system near the critical value \bar{t} are continuous although we cannot observe these statistics after \bar{t} as the GFP auction was replaced. In other words, bidders would have continued to behave as they would have before the auction rule change as if there had been no auction switch. Therefore, any bidding behavior change will be attributed to the treatment, or the launch of the new auction. Instead of making ad hoc assumptions about the bidding behaviors long before and after \bar{t} and running regressions across all the observations as the OLS method does, the RDD estimation uses nonparametric regressions to only estimate the bidding behaviors shortly before and after \bar{t} , and then compares the behavior difference around \bar{t} to achieve identification.

Because the above continuity assumption is addressing local properties, the estimated result also only reveals the local properties. It is worthwhile pointing out that Regression Discontinuity approach does not identify the average treatment effect $E[\alpha_i(x_i, t)]$ of the whole population. Instead, it identifies the local average $E[\alpha_i(x_i, t)]$ at \bar{t} , which is $E[\alpha_i(x_i, t)|t = \bar{t}]$.

5.2 RDD Estimator

Imbens and Lemieux (2007) and Van der Klaauw (2007) have very good surveys for the literature of RDD, especially the estimation methods used in RDD. The asymptotical boundary properties of the standard kernel estimator are not ideal because of the poor convergence rate, as pointed out by Hahn et al. (2001) and Porter (2003). Therefore, we consider the local linear regression method proposed by Fan and Gijbels (1996).

Let α_y and β_y solve the following minimization problems for the numerator:

$$\min_{\alpha_{yl}, \beta_{yl}, \alpha_{yr}, \beta_{yr}} \sum_{i|\bar{t}-h < t_i < \bar{t}} (y_i - \alpha_{yl} - \beta_{yl} \cdot (t_i - \bar{t}) - \delta \cdot X_i)^2 + \sum_{i|\bar{t} < t_i < \bar{t}+h} (y_i - \alpha_{yr} - \beta_{yr} \cdot (t_i - \bar{t}) - \delta \cdot X_i)^2$$

Here h is the bandwidth on either side of the discontinuity point. X_i is the covariate vector in which the estimation includes the market dummy variable and the weekday dummy variable.

Then the estimator for the average treatment effect will be $\hat{\tau} = \hat{\alpha}_{yr} - \hat{\alpha}_{yl}$.

6. Results

To estimate the bidding behavior difference, we try two different bandwidths on 11 days and 6 days around the discontinuity point on June 26, 2002. In the literature, there are some well-known methods for choosing this bandwidth optimally. Imbens and Lemieux (2007) have very good surveys on the optimal bandwidth selection. Due to the length of data for RDD, which is around 15 days before and after the cutoff point, we pick these two bandwidths. And readers should be noted that the results are not affected by the picks of the bandwidths. The bidding statistics examined here include the daily bid changing frequency, the daily bid range, the daily maximum bid, the 75th percentile bid, the mean bid, the median bid, the 25th percentile bid, and the daily individual bidder payment. For the last statistic, because we cannot observe the click-through-rate of each bidder's link, we simply assume every bidder received one unit of clicks every 15 minutes.

Table 5 shows the RDD estimation results, which are consistent with the OLS regression results shown in Table 4. Column 1 shows the frequency of individual daily bidding increased 3.7 times, which represents a 19% increase. For each individual bidder, bid range also increased by 71 cents, or 46% in relative value. All reject the hypothesis made by the static GSP auction framework

The estimation results also present how an individual bidder's payment changed each day after the new auction was launched. Table 5 shows that each bidder's average daily payment decreased by about 11.7%. One of the reasons that this number is so big might be that we cannot observe the actually click through rates on each slot and have to calibrate the numbers from Brooks (2005). This might bring bias to the magnitude of the estimation.

Table 5 also presents how an individual bidder's bid distribution changed after the launch of the GSP auction. Column 2 shows that the maximum bid and the 75th percentile bid tended to increase and the mean bid, median bid and the 25th percentile bid tended to decrease, which is also consistent with the increase of bid range.

Lastly, Figure 3 plots how each of the statistics evolves before and after June 26, 2002. We fit the estimation results with a smooth function.

When Yahoo! launched the new auction mechanism, it named the new bid "maximum willingness to pay", hoping bidders would simply bid their highest possible payment. Yahoo! hoped this would reduce the instability of the system and increase its revenue. But the above results suggest these goals were not achieved. Instead of reducing the strategic behaviors, bidders submitted their bids in a bigger range and changed their bids at a higher frequency.

6.1 Robustness Check

The above estimation provides the average effect across the markets. We also conduct the RDD estimation of the individual daily bidding frequency change and daily bid range for each market m . Figure 4 provides the histograms of the RDD estimation results. Most of daily bidding frequency increase resides between 0 and 10 times. The mean and median are also between 0 and 10. Meanwhile, the histogram of the daily bid range increase shows that in most of the markets bidders' increase their daily bid ranges and the mean and median value are around \$1. All these results reject the hypothesis and suggest the robustness of the above RDD estimation regarding the daily bidding frequency increase and the daily bid range increase.

6.2 Learning

One possible factor, which might have an impact on the above estimation, is the learning curve. If bidders were testing and learning the new auction system, the estimation of the bidding frequency difference may have biases.

There is no good control for learning because of the limitation of the data, especially to what extent learning may affect estimation. However, one important fact is that, before Yahoo! launched the GSP auction in June, Google had already started a similar GSP auction in April (Jansen and Mullen, 2008). Serious advertisers would have usually had accounts with both of the search engines' sponsored search auctions. Therefore, ideally, the bidding behaviors under the Google GSP auction would be a perfect control group for the current RDD analysis. Unfortunately, it is even harder to obtain keyword auction data from both of these two search engines. However, as keyword advertising is a serious business that deals with a serious amount of money, it is reasonable to assume that most of the bidders in the Yahoo! GSP auction had already experienced and learned the GSP auction bidding under Google. Therefore, when they faced the GSP auction under Yahoo!, learning might not play an important role.

7. Conclusion

The sponsored search auction is an important and interesting mechanism worth great attention. Edelman et al. (2007) and Varian (2007) provide a foundation to understand it. However, the actual bidding in the real sponsored search auction is much more complicated, especially when a significant portion of bids are submitted by automated bidding robots following certain mechanical bidding rules. When one bidder can change his bid over 10 000 times in one a day, an immediate question should be whether the current theory has fully captured the main features of the actual auction and is able to give accurate predictions about bidding behavior.

Fortunately, the institutional variation of the Yahoo! sponsored search auction rule provides an opportunity to examine the theory using a huge data set with over 1 million observations. This paper provides solid evidence suggesting that the GSP auction framework may provide biased predictions on the actual bidding.

Due to the limitations of the data set, it is not the goal of this paper to interpret why the actual bidding does not match the current GSP auction framework predictions. But, we conjecture that the automated bidding in the sponsored search auction may affect the explaining power of the current theory framework. An interesting topic would be to examine what role automated bidding behaviors play in the actual bidding process. This will be left for future research.

References

- Angrist, J., & Lavy, V. (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement, *The Quarterly Journal of Economics*, 114, 533–75. <http://dx.doi.org/10.1162/003355399556061>
- Athey, S., & Ellison, G. (2007). Position auctions with consumer search, *Working Paper*.
- Borgers, T., Ingemar C., Martin P., & Vaclav P. (2007). Equilibrium bids in sponsored search auctions: theory and evidence, *Working Paper*.
- Brooks, N. (2005). The atlas rank report: How search engine rank impacts traffic, *Digital Marketing Insights*
- Cameron, C., & Pravin T. (2005). *Microeconometrics: methods and applications*, Cambridge University Press.
- Cary, M., Aparna D., Benjamin E., Ioannis G., Kurtis H., Anna R., Claire M., & Mickael S. (2008). On best-response bidding in GSP auctions, *NBER Working Paper*.
- Chen, S., & Wilbert K. (2008). The work disincentive effects of the disability insurance program in the 1990s. *Journal of Econometrics*, 142, 757–84. <http://dx.doi.org/10.1016/j.jeconom.2007.05.016>
- Van der Klaauw, W. (2002). Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach. *International Economic Review*, 4, 1249–87. <http://dx.doi.org/10.1111/1468-2354.t01-1-00055>
- Van der Klaauw, W. (2007). Regression-discontinuity analysis. *The New Palgrave Dictionary of Economics*.
- Edelman, B., & Michael O. (2006). Strategic bidder behavior in sponsored search auctions, *Decision Support Systems*
- Edelman, B., Michael O., & Michael S. (2007). Internet Advertising and the generalized second-price auction: selling billions of dollars worth of keywords. *American Economic Review*, 1, 242–59. <http://dx.doi.org/10.1257/aer.97.1.242>
- Fan, J., & Irene G. (1996). *Local polynomial modelling and its applications*, Chapman and Hall, London.
- Hahn, J., Petra T., & Wilber Van der K. (2001). Identification and estimation of treatment effects with a regression-discontinuity design, *Econometrica*, 1, 201–9. <http://dx.doi.org/10.1111/1468-0262.00183>
- Imbens, G., & Joshua A. (1994). Identification and estimation of local average treatment effects, *Econometrica*, 2, 467–75.
- Imbens, G., Joshua A., & Thomas L. (2007). Regression discontinuity designs: a guide to practice, *NBER Working Paper*, 13039. <http://dx.doi.org/10.2307/2951620>
- Jansen, B. J., & Tracy M. (2008). Sponsored search: an overview of the concept, history, and technology. *Int. J. Electronic Business*, 2, 114–31. <http://dx.doi.org/10.1504/IJEB.2008.018068>
- Lee, D. (2008). Randomized experiments from non-random selection in US house election. *Journal of Econometrics*, 142, 675–97. <http://dx.doi.org/10.1016/j.jeconom.2007.05.004>
- McAdams D., & Michael S. (2007). Who pays when auction rules are bent?. *International Journal of Industrial Organization*, 5, 1144–57. <http://dx.doi.org/10.1016/j.ijindorg.2006.09.004>
- Porter, J. (2003). Estimation in the regression discontinuity model, Working Paper, Department of Economics, University of Wisconsin.
- Varian Hal, R. (2007). Position auctions. *International Journal of Industrial Organization*, 6, 1163–78.
- Yuan, J. (2009). Backtesting advertisers automated bidding strategies, University of Minnesota, *Working Paper*.
- Zhang, X. (2005). Finding edgeworth cycles in online advertising auctions, *Working Paper*.

Table 1. Bid Statistics of the Top 10 Most Clicked Markets

Market	Observations	mean	stddev	min	Max
All 812 Markets	18,634,347	5.7	9.1	0.05	9,170
1	1,455,161	16.73	6.52	0.05	60
2	2,041,397	12.24	5.27	0.05	50
3	58,269	7.07	2.41	0.05	22.01
4	14,467	18.86	9.88	0.05	50
5	294,538	14.86	4.06	0.05	50
6	22,884	4.92	2.95	0.05	50
7	20,659	17.98	6.88	0.05	50
8	21,136	14.34	7.06	0.05	50
9	28,695	5.06	3.82	0.05	50
10	17,850	18.63	8.91	0.05	50

Table 2. The Average Bidding Frequency on Each Day

	Mean	Sd.	Max	Min
	14.9	128	17.867	1
Threshold	Portion of Automated Bids			
> 40	53.30%			
> 100	44.60%			
> 200	38.80%			
> 500	28.10%			

Note: N = 18,634,347.

Table 3. Summary Statistics from June 15th to July 5th

Before June 25th	Mean	Stv	Min	Max
Bid Frequency	20.9	108	1	4,934
Bid Range	0.617	1.55	0	48.99
Max Bid	2.66	3.69	0.05	100
75 percentile bid	2.74	3.89	0.05	100
Median Bid	2.54	3.59	0.05	100
Mean Bid	2.52	3.55	0.05	100
25 Percentile bid	2.38	3.48	0.05	100
After June 25th	Mean	Stv	Min	Max
Bid Frequency	23.3	143	1	6,011
Bid Range	0.983	3.33	0	49.95
Max Bid	3.02	4.64	0.05	100
75 percentile bid	2.82	4.1	0.05	50
Median Bid	2.59	3.75	0.05	50
Mean Bid	2.56	3.62	0.05	42.8
25 Percentile bid	2.3	3.42	0.05	50

Table 4. The Change of the Statistics in Three Months by OLS Regression

	α_1	α_2	α_3	R2
Bid Frequency	14.50%	28%	18.80%	0.28
	(-0.008)	(-0.009)	(-0.01)	
Bid Range	14.00%	14.40%	15.20%	0.26
	(-1.3%)	(-1.3%)	(-1.4%)	
Max Bid	10.30%	12.70%	15.10%	0.45
	(-0.7%)	(-0.7%)	(-0.8%)	
75 percentile Bid	9.70%	12.10%	14.60%	0.45
	(-0.7%)	(-0.7%)	(-0.8%)	
Mean Bid	9.10%	11%	13.60%	0.46
	(-0.007)	(-0.007)	(-0.008)	
Median Bid	8.80%	10.80%	13.70%	0.45
	(-0.007)	(-0.007)	(-0.008)	
25 percentile Bid	6.70%	8.20%	11.10%	0.42
	(-0.007)	(-0.008)	(-0.008)	

Table 5. RDD Estimation Results

	h=11		h=6	
	Absolute	Relative	Absolute	Relative
Bid Frequency	3.671	19.40%	3.11	26%
	(-2.031)	(-0.021)	(-2.284)	(-0.024)
Bid Range	0.709	46.00%	0.91	54%
	(-0.006)	(-0.006)	(-0.053)	(-0.039)
Daily Payment	-5.4	-11.70%	-11.28	-12.9%
	(-0.62)	(-1.42)	(-0.25)	(-0.03)
Max Bid	0.298	3.30%	0.43	9%
	(-0.051)	(-0.017)	(-0.066)	(-0.022)
75 percentile bid	0.049	0.00%	0.119	5%
	(-0.046)	(-0.017)	(-0.061)	(-0.019)
Median Bid	-0.113	-3.30%	-0.084	0.5%
	(-0.042)	(-0.017)	(-0.048)	(-0.02)
Mean Bid	-0.106	-2.20%	-0.078	-1.9%
	(-0.04)	(-0.017)	(-0.047)	(-0.019)
25 Percentile bid	-0.287	-9.70%	-0.309	-7.1%
	(-0.039)	(-0.019)	(-0.05)	(-0.021)

Notes: There are 1,099,781 bids collected from 812 keyword markets. "Absolute" measures the absolute value change; "relative" measures the percentage change. h is the bandwidth, taking values of 11 days and 6 days respectively. To analyze payment, we normalize the click on the first slot to one. The click declining rate follows Brooks (2005). Therefore, the relative change for the daily payment is more meaningful.



Search bar with 'refinance' entered and a 'Search' button.

Search Results

1 - 10 of about 40,100,000 for refinance - 0.42 sec. (About This Page)

Also try: [refinance mortgage](#), [auto refinance](#), [home refinance](#) More...

- **Refinance - Lendingtree**
www.LendingTree.com - Refinance \$200,000 for \$667/Month. Refinance Offers Quality Online.
 - **Bad Credit? Refi Today**
www.FullSpectrumLending.com - Homeowner in debt? Need cash now? Fast home refi. Low payments.
 - **Nationpoint Home Loans**
www.nationpoint.com - First time buyer specialists. 0% down loans with credit scores 620+
 - **Refinance**
OriginsRateSlide.com - Rates Still Near Historic Lows. Get and Compare Your Rates Now.
1. **Mortgage and Loan Interest Rates at Yahoo! Real Estate**
Find up to date national mortgage interest rates for fixed rate mortgages, ARM adjustable rate mortgages and interest only mortgages at Yahoo! Real Estate
Quick Links: [Mortgage Calculators](#) - [Mortgage Rates](#)
realestate.yahoo.com/loans - [More from this site](#)
 2. **Mortgage Refinance and Home Equity Loans - GetSmart.com**
A service of LendingTree - Complete a short 2-minute form & get up to 5 free mortgage quotes - no obligation and no Social Security Number required. Bad Credit OK.
Quick Links: [Home Refinancing](#) - [Fixed Home Equity](#) - [Mortgage Quotes](#)
www.getsmart.com - [More from this site](#)
 3. **Quicken Loans - Home loans, Refinancing, Interest-only options**
Quicken Loans - Get information, check rates, and learn about refinancing your current home loan. Compare mortgage options, apply online, get pre approved and close fast.
Quick Links: [Get Mortgage Rates](#) - [Compare Home Loans](#) - [Interest-Only Loans](#)
www.quickenloans.com - [More from this site](#)
 4. **Real Estate Financing in the Yahoo! Directory**
Browse through a long list of companies that offer home loans, refinancing, and ... Offers home refinance, secured debt consolidation, and other lending services. ...
dir.yahoo.com/Business_and_Economy/.../Real_Estate/Financing - 16k - [Cached](#) - [More from this site](#)
 5. **Mortgage: Quotes, Rates, Loans & Refinance by National Mortgage**
Free mortgage quotes, calculators and guides, with unmatched customer service. ... Mortgages free mortgage quotes Home Loans Refinance credit calculator house loan ...
Category: [Real Estate Financing](#) > [Brokers/agents](#)
www.nationalmortgage.com - [More from this site](#)
 6. **E-Loan: Mortgage, Refinance, Home Equity, Auto Loans, Savings, CDs**
E-Loan offers home mortgage, refinance, home equity loans, lines of credit, auto and motorcycle loans, savings accounts and CDs with great rates online. No hidden fees.
Quick Links: [Home Equity](#) - [Mortgage Refinance](#) - [Savings Accounts](#)
www.eloan.com - [More from this site](#)
 7. **Mortgage Refinance and Home Loans - Ameriquest**
Choose AmeriquestMortgage.com to find a great mortgage. Ameriquest provides home mortgage loans, mortgage refinancing, and debt consolidation services.
Category: [Real Estate Financing](#)
www.ameriquetmortgage.com - [More from this site](#)
 8. **Home Loans - Equity Refinance, Mortgage & Auto | LendingTree**
LendingTree - Your ARM could increase as much as 60%. Lock in a low rate now. \$175,000 for \$900/mo. 15 year Fixed. 1 simple form, get up to 4 custom offers in minutes.
Quick Links: [Refinance Now](#) - [Fixed Home Equity](#) - [Mortgage Quotes](#)
www.lendingtree.com - [More from this site](#)
 9. **Refinance Articles - Quicken Loans**
Refinance Articles - Quicken Loans offers mortgages, home loans, refinance and home equity loans. Find information to help you make informed mortgage decisions.
www.quickenloans.com/refinance/articles/index.html - [More from this site](#)
 10. **Refinance Home Mortgage - Ditech.com**
Save on your monthly payments or use your equity to get cash out with the Ditech's home mortgage refinance products.
www.ditech.com/refinance/index.html - [More from this site](#)
- **Countrywide@ Home Loans**
www.Countrywide.com - Fast home refi, good credit or not. Countrywide@. 4 out of 5 approved.
 - **Get up to \$1,500 - Fast and Easy**
payday-fastloan.com - No fax.No check.Payday loan Personal loan And many more.

SPONSOR RESULTS

- SPONSOR RESULTS
- Countrywide@ Home Loans**
Fast home refi, good credit or not. Countrywide@. 4 out of 5 approved.
www.Countrywide.com
 - Get up to \$1,500 - Fast and Easy**
No fax.No check.Payday loan Personal loan And many more.
payday-fastloan.com
 - Refinance**
Get the most out of your mortgage. Refinance with GMAC Mortgage.
www.gmacmortgage.com
 - Refinance with Ditech@**
Get Low Fixed Rates, Lending Costs Quick Approval. Apply Online Now.
www.ditech.com
 - Wachovia Pick-a-Payment**
Lower Payments, Increase Cash Flow. New from Wachovia. Learn More.
www.Wachovia.com
 - Refinance**
Rates Still Near Historic Lows. No Lender Fee. Approval in Minutes.
www.eloan.com/refinance
 - Compare Refinance Quotes**
Complete Our Easy Form & Receive Up To 4 Low Refinance Quotes.
www.GuideToLenders.com/refinanc
 - Refinance Quotes - Save \$1000s Now**
Refinance your mortgage loan before rates explode. Get matched here.
www.usloanquotes.com
 - [See your message here...](#)

SPONSOR RESULTS

Also try: [refinance mortgage](#), [auto refinance](#), [home refinance](#) More...

1 2 3 4 5 6 7 8 9 10 Next

Figure 1.Sponsored Links for the Keyword "Refinance"

Keyword: **Keywords Dynamo**
 Welcome to the NEW Keyword Dynamo Tool. The only Tool where you can view Overture Bids for US Market. Due to it popularity and usage overture has blocked

Overture Keyword Suggestion
 Overture Bids

View Bids

Type in a search term and we'll show you the Max Bids and listings for that term.

- 1. Refinance - Lendingtree**
\$300,000 For Only \$1,000/month. Refinance Today. Bad Credit Options.
www.LendingTree.com
(Advertiser's Max Bid: \$16.13)
- 2. Bad Credit? Refi Today**
Homeowner in debt? Need cash now? Fast home refi. Low payments.
www.FullSpectrumLending.com
(Advertiser's Max Bid: \$12.53)
- 3. Nationpoint Home Loans**
First time buyer specialists. 0% down loans with credit scores of 620+.
www.nationpoint.com
(Advertiser's Max Bid: \$9.70)
- 4. Refinance**
Rates Still Near Historic Lows. Get and Compare Your Rates Now.
Oregon.RateSlide.com
(Advertiser's Max Bid: \$9.55)
- 5. Countrywide@ Home Loans**
Refi to combine 1st mortgage & debt. Low payments with a 40-year loan.
www.Countrywide.com
(Advertiser's Max Bid: \$8.75)
- 6. Get up to \$1,500 - Fast and Easy**
No fax.No check.Payday loan.Personal loan.And many more.
payday-fastloan.com
(Advertiser's Max Bid: \$8.42)
- 7. Refinance**
Refinance and reduce your payments upto 60%.
www.floridaslowestrates.com
(Advertiser's Max Bid: \$7.49)
- 8. Refinance**
Get the most out of your mortgage. Refinance with GMAC Mortgage.
www.gmacmortgage.com

<http://keyword.secretstohighprofit.com/default.aspx>

3/28/2007

Figure 2.Bids and Placements

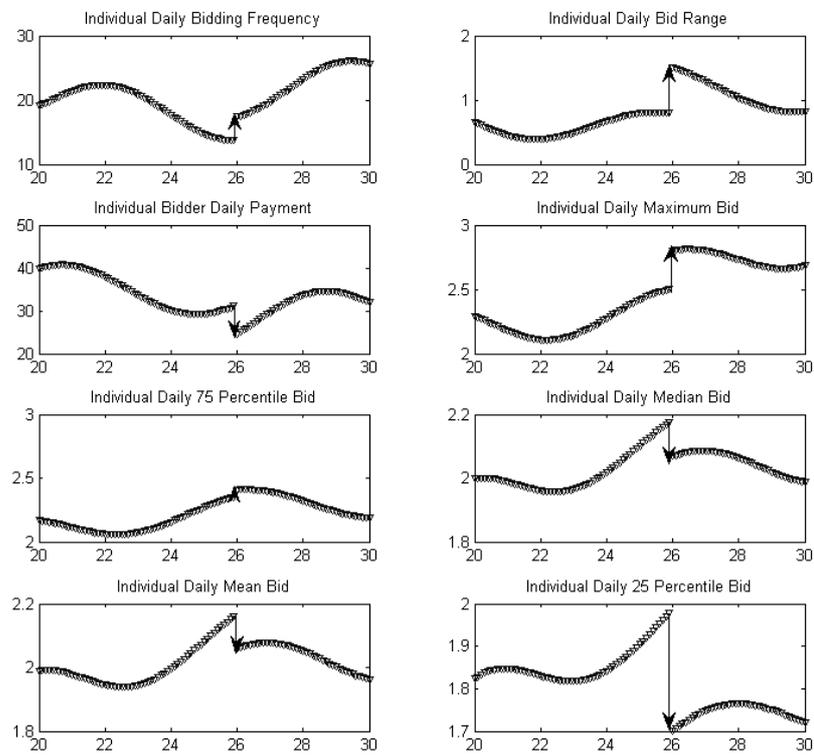


Figure 3. The Impact of GSP Auction on Bidders' Behavior

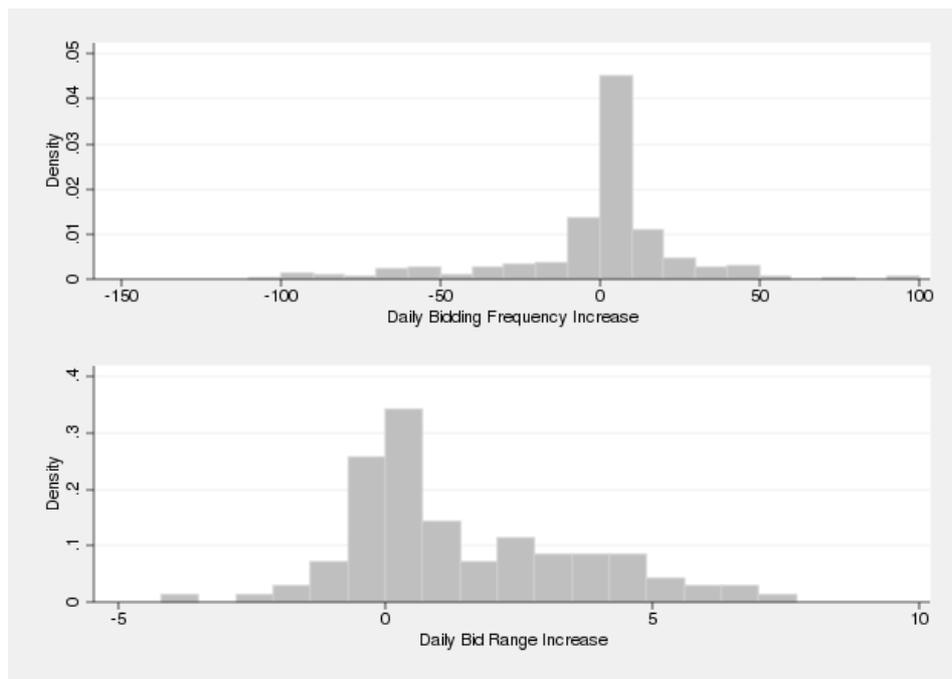


Figure 4. Robustness Check on the Histograms of the Daily Bidding Frequency Increase and the Daily Bid Range Increase across Markets